Nucleic Acids Research

Nucleic Acids Res. 2024 Nov 18;53(D1):D1697–D1710. doi: 10.1093/nar/gkae1116

NASA open science data repository: open science for life in space

Samrawit G Gebre [1,2], Ryan T Scott [2,2], Amanda M Saravia-Butler [3], Danielle K Lopez [4], Lauren M Sanders [5], Sylvain V Costes [6,✉]

Author information Article notes Copyright and License information

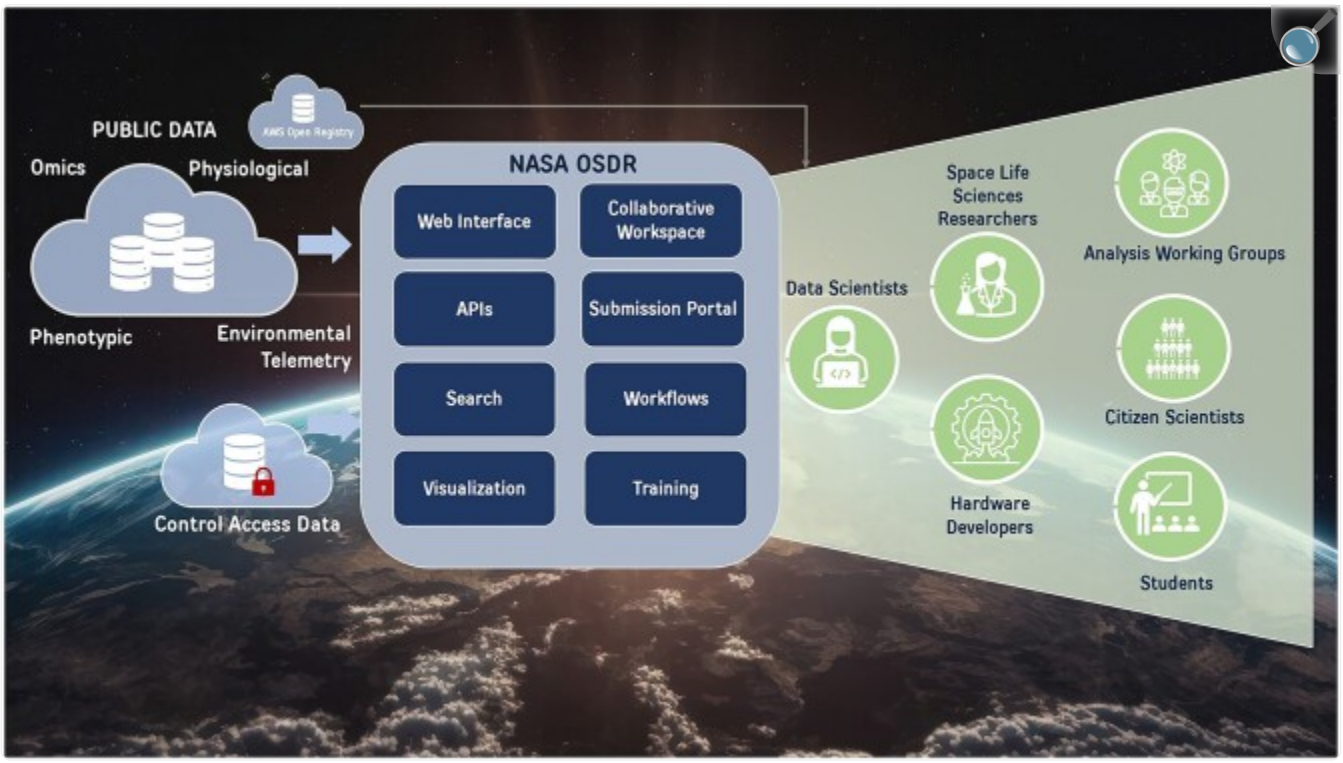
PMCID: PMC11701653 PMID: 39558178

Abstract

Space biology and health data are critical for the success of deep space missions and sustainable human presence off-world. At the core of effectively managing biomedical risks is the commitment to open science principles, which ensure that data are findable, accessible, interoperable, reusable, reproducible and maximally open. The 2021 integration of the Ames Life Sciences Data Archive with GeneLab to establish the NASA Open Science Data Repository significantly enhanced access to a wide range of life sciences, biomedical-clinical and mission telemetry data alongside existing ‘omics data from GeneLab. This paper describes the new database, its architecture and new data streams supporting diverse data types and enhancing data submission, retrieval and analysis. Features include the biological data management environment for improved data submission, a new user interface, controlled data access, an enhanced API and comprehensive public visualization tools for environmental telemetry, radiation dosimetry data and ‘omics analyses. By fostering global collaboration through its analysis working groups and training programs, the open science data repository promotes widespread engagement in space biology, ensuring transparency and inclusivity in research. It supports the global scientific community in advancing our understanding of spaceflight's impact on biological systems, ensuring humans will thrive in future deep space missions.

Graphical Abstract

Graphical Abstract.



Open in a new tab

Introduction

The fields of space biology and space health are rapidly expanding and becoming more essential as humanity embarks on missions to the Moon, Mars, and other deep space missions beyond low Earth orbit. The heightened biomedical risks and new challenges associated with deep space missions (1) necessitate the discovery of new knowledge, the development of new health countermeasures and the creation of novel ecosystems (2,3), life support systems, crop production methods, *in situ* biomedical research and biomedical support capabilities (4). Supporting distant and long-duration missions is our new paradigm that requires biological and health-related data to be findable, accessible, interoperable and reusable (FAIR) (5,6). Optimal data governance for public dissemination encompasses a spectrum from closed to mediated, embargoed and fully open access. Scientific reproducibility, knowledge discovery and data reuse is maximized when data are both FAIR and open (7).

For the past decade, NASA GeneLab ‘omics database has played the leading role in collecting, curating, processing and organizing space biology data (8), catalyzing numerous biological discoveries. Noteworthy publications that have utilized GeneLab include studies on mitochondrial dysfunction across species (9), short duration spaceflight on human health (10), kidney dysfunction (11), endocrinology (12), carbon dioxide impact (13), skin health (14), liver lipid dysregulation (15) and changes in the gut microbiome (16). Such pioneering work highlighted the need for all space biomedical data beyond molecular ‘omics to also be maximally open access and FAIR. Consequently in 2021, the integration (17) of the NASA Ames Life Sciences Data Archive (ALSDA) with the GeneLab database culminated in the creation of the NASA Open Science Data Repository (OSDR; https://www.nasa.gov/osdr/). Today, OSDR combines ‘omics with non-omics life sciences space-related data that span physiological, phenotypic, bioimaging, behavioral, clinical, environmental telemetry, hardware data and more. This integration involved ALSDA transforming to standardize, curate and make all its collected data more FAIR and open (7,17).

This paper describes the new OSDR database architecture and features that enable the submission, hosting, retrieval and analysis of diverse data types, including both ‘omics and non-omics, from spaceflight to analog missions. We review the new biological data management environment (BDME), which supports the submission of both ‘omics and non-omics data in a standardized format. We introduce the new OSDR database layout, which includes access to additional experimental, payload and mission-specific information and telemetry data associated with each study. We also describe the new controlled data access feature that mediates the sharing of sensitive commercial astronaut data, making data as open as possible and as close as necessary. The generation of novel spaceflight data by the GeneLab Sample Processing Laboratory (SPL) is described, along with the expansion of GeneLab's standardized data processing pipelines and available processed data products. Access to all metadata and data hosted on OSDR is provided by the expanded Application Programming Interface (API) and Registry of Open Data on Amazon Web Services (AWS). We also review the rich set of visualization tools available to easily analyze and interpret OSDR data, further democratizing these data. We also present user roadmap examples for the public to engage in spaceflight biomedical sciences leveraging OSDR. Lastly, we introduce the OSDR Analysis Working Groups (AWGs) and training initiatives that enhance engagement with space-relevant data and support a robust and active space biology and space health research community.

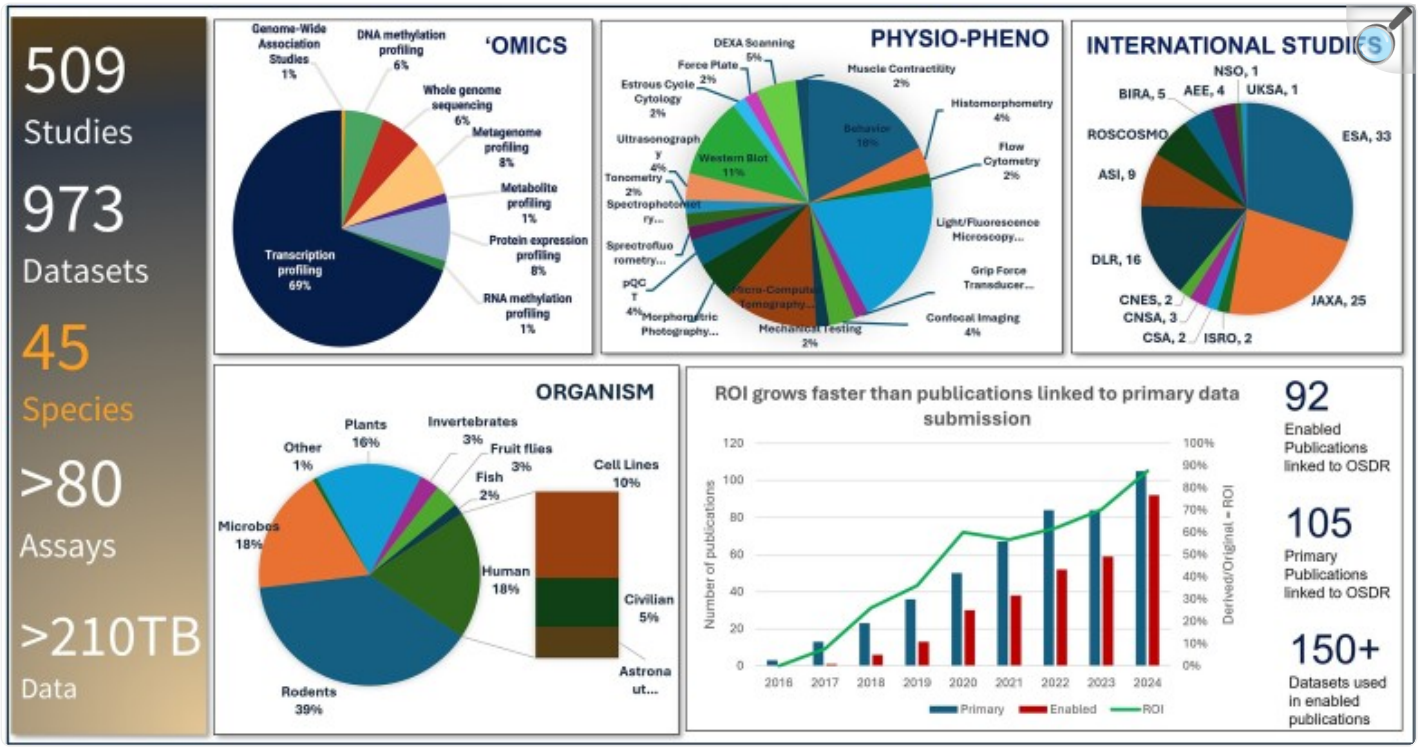
Expanded database architecture and new features

Database content and use

OSDR champions data stewardship to enable all humanity to participate in discovery and exploration by making data more equitable and accessible to all (18). As of October 2024, OSDR houses over 500 studies, including close to 1000 spaceflight or space-analog datasets, with data from over 80 different assays, encompassing both ‘omics and physiological/phenotypic assays (Figure 1). OSDR is species-agnostic, offering access to a diverse range of studies from rodents to plants, microbes and humans—including data from commercial astronauts connected to the Inspiration 4 SpaceX mission (10). The data hosted on OSDR are linked to 105 original peer-reviewed scientific publications and

have enabled 92 publications, representing an 88% return on investment (ROI), measured by the number of enabled publications versus primary publications (Figure 1). This indicates that the ROI grows faster than the number of primary publications because the data in OSDR have led to a significantly higher number of new publications, suggesting that each dataset submitted to OSDR has the potential to generate multiple follow-up studies and accelerate research output. The repository is also currently curating data from other space agencies and companies such as the European Space Agency (ESA), providing studies on human head-down bed rest, human dry immersion, model organisms exposed to simulated space radiation and other model organisms flown on the International Space Station (ISS). The demand for OSDR to ingest commercial human data has been surging, highlighting its unique role in the space biology ecosystem. Not only does OSDR promote greater FAIR/Open access to data, but it also enables comparisons across hundreds of datasets, spanning different species and various missions/experiments. This capability greatly enhances the scientific value of space exploration by enriching the biomedical insights gained from shared data.

Figure 1.



Open in a new tab

OSDR metrics. Overview of studies hosted on OSDR as of October 2024, including the total number of international studies, percent of total datasets for each ‘omics and each phenotypic-physiological measurement type, percent of total studies for each organism and the ROI measured by the number of enabled publications versus primary publications that use OSDR datasets.

Data management workflow

To increase the reusability and accessibility of biological data from space-relevant experiments and data types and to improve data submission planning, the original GeneLab submission portal was expanded into the BDME. Key features such as Digital Object Identifiers (DOIs) and the ISA framework (Investigation, Study, Assay) (19,20) remain integral to the BDME. The submission portal now has the capability to collect all types of data-metadata, encompassing not just ‘omics but all life science and biomedical assays. The enhanced BDME tools are designed to ease the submission process, enabling self-curation with established metadata and data standards through a new user-friendly interface and a streamlined workflow.

The BDME's latest updates support ongoing data and metadata submission throughout research, avoiding end-of-grant rushes that compromise data quality, while preserving data privacy. BDME introduces a collaborative workspace, enabling users to privately share, organize and store files during their experiments. The collaborative workspace provides users a new solution for transferring large data files quickly and securely, with the File Application for Seamless Transfer to/from the Workspace (FASTWork). FASTWork ensures the integrity of data transfers by automatically resuming the process if interrupted due to connectivity issues or computer downtime. This functionality is accessible through both a command-line interface and a downloadable application, making it versatile for various user preferences and technical environments. Users can also add collaborators and specify permissions and roles such as editors, supervisors and submitters, enhancing the collaborative and efficient management of research data. Additionally, a private link feature lets researchers securely share dataset previews with peer reviewers and editors.

The Research Data Submission Agreement (RDSA), now a requirement under the NASA Biological and Physical Sciences (BPS) Division's Scientific Data Management Policy (https://science.nasa.gov/biological-physical/data/), guides NASA investigators in outlining data products, assays, tissues and submission timelines. The information captured in the RDSA generates a unique Experiment page, automatically populating the data to reduce further input from principal investigators (PIs), thus significantly lowering the burden of data entry and ensuring metadata richness—a key FAIR requirement. Importantly, aside from RDSA-directed NASA BPS policy timelines, there is no set timeframe for data release from submission to public availability on OSDR, providing submitters control over their data's accessibility. Since there is no required timeframe for public release after data submission, submitters are encouraged to begin the submission process as early as possible. Starting early helps researchers meticulously collect and standardize their metadata throughout their experiments and prevents a heavy data submission workload at the end of their grant. This ensures ample time to gather complete metadata, resulting in higher quality and more comprehensive curation records, ultimately streamlining the submission process.

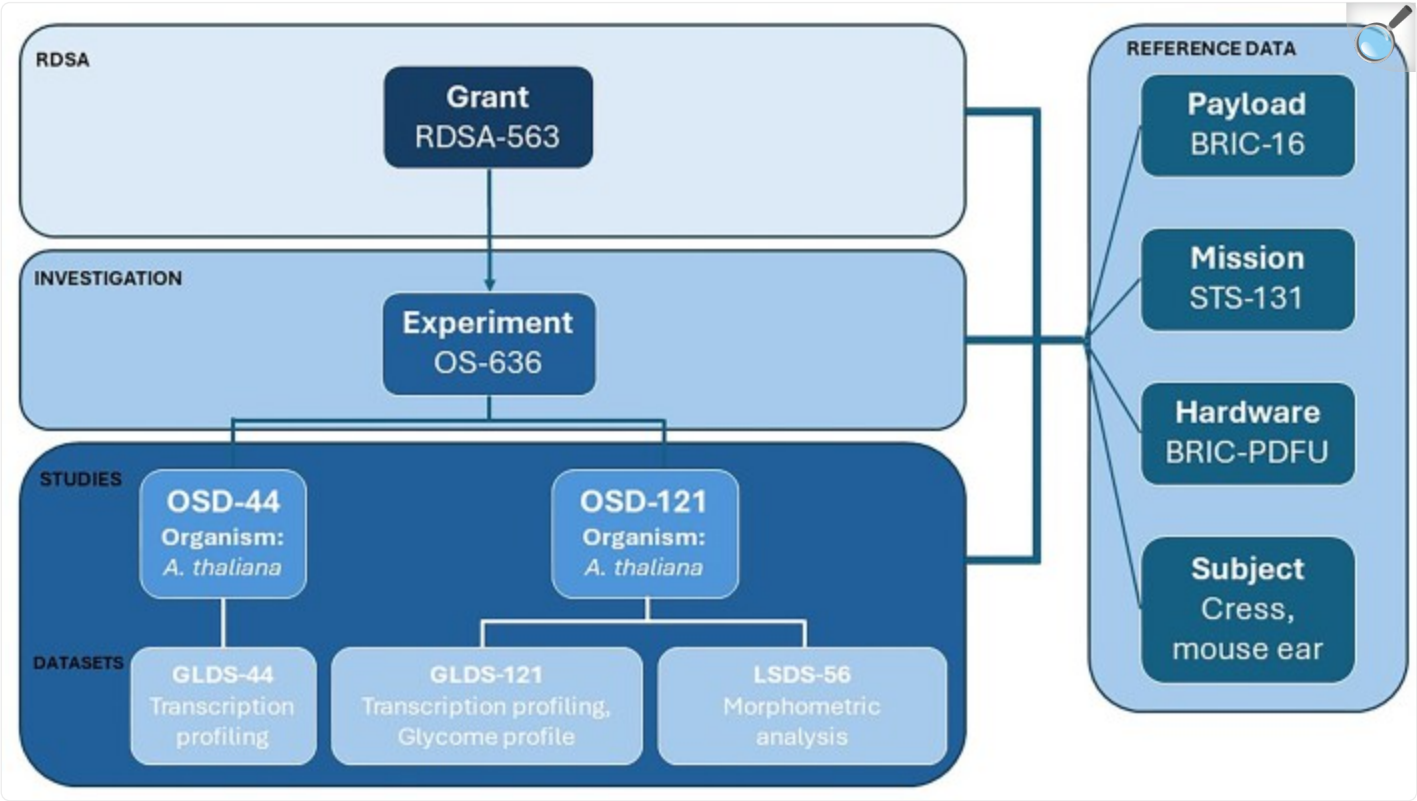
New data collections

Integrating ALSDA data into OSDR required developing new data schemas to manage diverse records such as experiments, biospecimens, hardware, payload and mission data. To achieve this, OSDR expanded on GeneLab's existing database to create a schema linking studies with their corresponding reference data collections.

To support the legacy ALSDA data structure used for the past 40 + years and meet new requirements, OSDR introduces a workflow based on a new data and metadata hierarchy (Figures 2 and 3). This new system offers two distinct routes for data submission, one for NASA-funded data (Figure 2), and one for non-NASA funded data (Figure 3). For NASA-funded PIs, the process begins with the RDSA, which is the initial step and *first tier* of submission. RDSAs capture comprehensive grant-level information, organism metadata and anticipated datasets for submission (Figure 2). The *second tier* in OSDR's data management workflow describes all Experiments performed with grant funds, and establishes connections between individual investigations, linking them to related study collection, and reference data

collections (Figure 2). The *third-tier* guides submitters through the data submission process, offering options for either a structured, guided submission or independent data and metadata entry for each dataset within a study (Figure 2). This tier is the point of entry for non-NASA PIs who do not need to have a NASA investigation identifier (Figure 3).

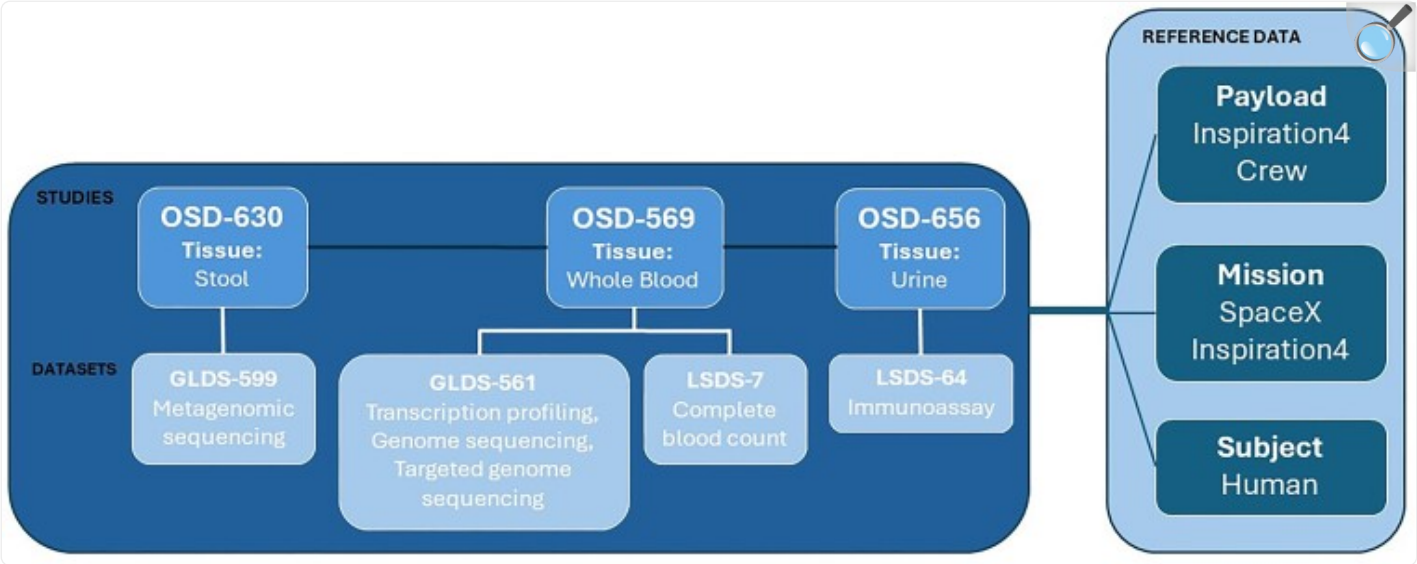
Figure 2.



Open in a new tab

NASA funded PI data and metadata hierarchy. The RDSA is for NASA-funded PIs and describes the experiment(s) conducted with the associated grant funds. Each experiment receives a unique identifier (e.g. OS-636) and is linked to respective Reference data and studies. Reference data records associated with each experiment include payload (e.g. BRIC-16), mission (e.g. STS-131), hardware (e.g. BRIC-PDFU) and subject (e.g. Cress, mouse ear). Each study associated with an experiment has a unique identifier (e.g. OSD-44 and OSD-121). Most studies are categorized based on organism and tissue type, although plant studies are distinguished based solely on organism as shown here. Each study is further categorized between ‘omics and phenotypic-physiological datasets. GLDS includes all ‘omics data and LSDS contains all phenotypic-physiological data derived from the respective study.

Figure 3.



Open in a new tab

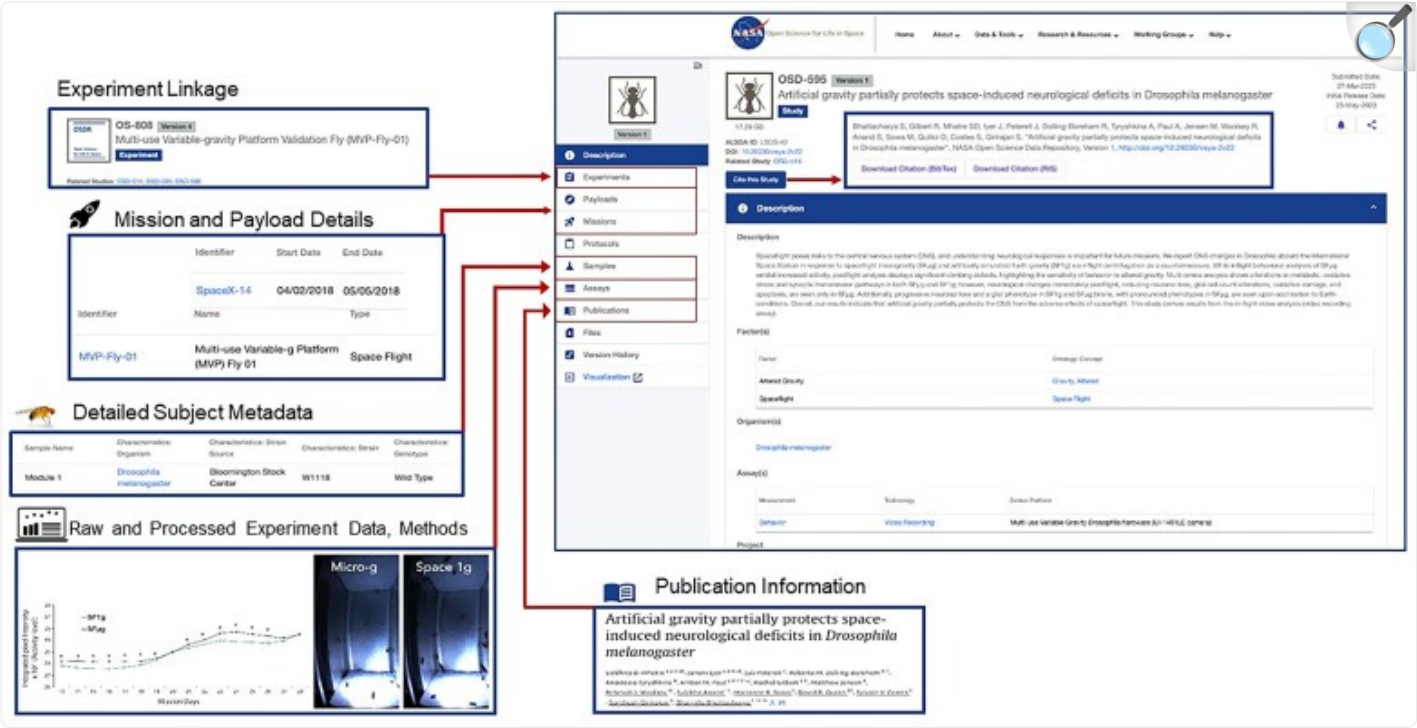
Non-NASA funded PI data and metadata hierarchy. The point of entry for non-NASA PIs begins at the study level. Most studies are categorized based on organism and tissue type, although plant studies are distinguished based solely on organism. Each study can maintain links to respective reference data when applicable. Reference data records associated with each study can include payload (e.g. Inspiration4 Crew), mission (e.g. SpaceX Inspiration4), subject (e.g. Human) and/or hardware (not included in this study). Each study has a unique identifier (e.g. OSD-630, OSD-569, OSD-656). Each study is further categorized between ‘omics and phenotypic-physiological datasets. GLDS includes all ‘omics data and LSDS contains all phenotypic-physiological data derived from the respective study.

As illustrated in Figure 2, OSDR utilizes persistent identifiers throughout the hierarchy of its data schema. Specifically, NASA investigators are assigned a unique identifier for each RDSA (e.g. RDSA-563) and experiment (e.g. OS-636; OS = Open Science experiment), while all data submitters receive a unique study identifier (e.g. OSD-44; OSD = Open Science Dataset). Within this framework, an experiment can encompass multiple studies. The majority of studies are categorized by organism and tissue type, whereas plant studies are categorized solely by organism, often encompassing multiple tissue types within a single study. Each study, in turn, may include up to two types of datasets: ‘omics assays (e.g. GLDS-44; GLDS = GeneLab Dataset) and/or physiological/phenotypic assays (e.g. LSDS-56; LSDS = Life Science Dataset) and all assays of the same type are grouped under a single unique dataset identifier. This hierarchical arrangement promotes thorough organization and enhances the ease of data management and retrieval, supporting

detailed scientific research, analysis, and the understanding of study design based on the original data.

These new collections are seamlessly integrated into the new OSDR repository webpage, making them findable using the OSDR repository search features, and they are also linked within each OSD study page, directly connecting studies to corresponding experiment, mission and payload pages (Figure 4). In addition to these reference data record links, each study page consists of descriptive and rich metadata, links to definitions of ontology terms and detailed sample and assay metadata that follow community standards. Users can also access all available data files structured by data level, publication information and author contact details from a study page (Figure 4). Additionally, the interface supports easy citation of datasets with a DOI, ensuring proper credit and discoverability.

Figure 4.



Open in a new tab

An OSD study page. Screenshot of study OSD-595 showing the links to the respective experiments (OS-808), mission (SpaceX-14) and payload (MVP-Fly-01) details, sample metadata, GLDS and LSDS assay metadata and primary publications (left) and study citation (inset).

The OSDR reference data collection is a new collection which reflects the integration of more than 2000 + records that

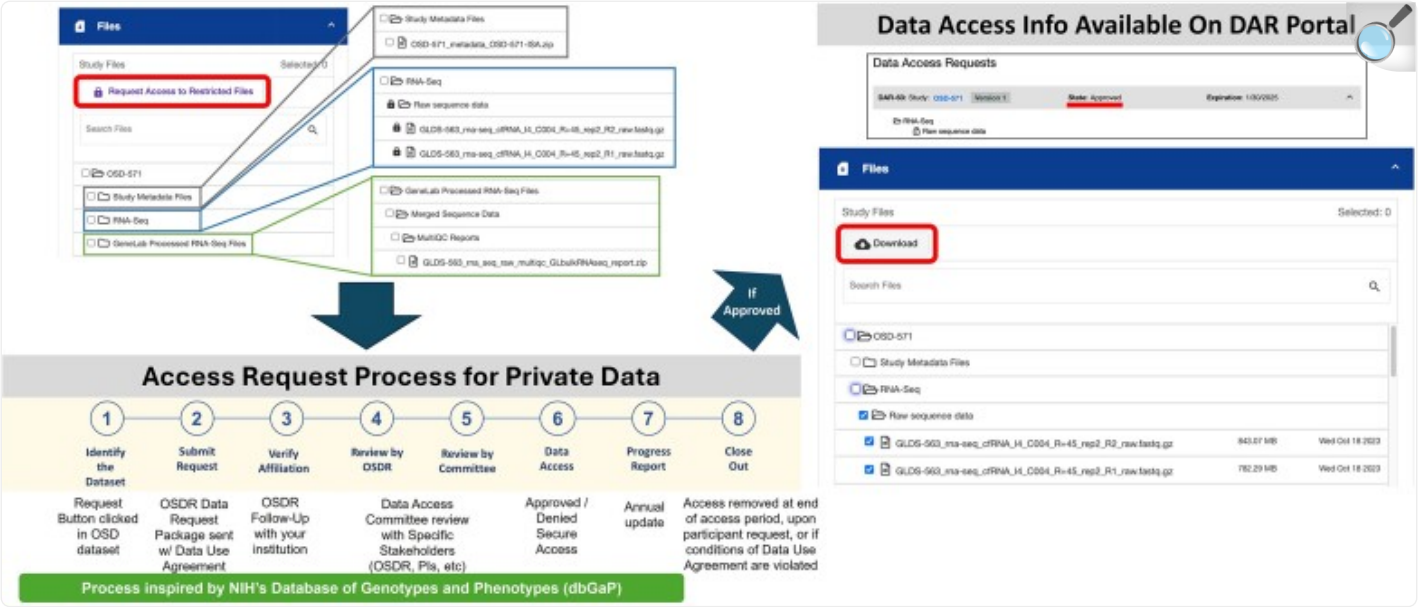
span over 40 years of life sciences research migrated from the former ALSDA legacy relational database. In addition, if submitters lack the necessary reference data, OSDR's curation team is available to assist in adding these to the repository and refining the collection. This flexibility ensures that non-NASA PIs can fully engage with the OSDR ecosystem without being hindered by NASA-specific administrative details.

Controlled data access

Although almost all data available in OSDR are openly accessible, sensitive human-derived data are secured through a controlled access system based on the NIH Database of Genotypes and Phenotypes (21). This system allows storage and sharing of data in accordance with stringent agreements and policies established by the Institutional Review Board (IRB) of the data submitter. By adhering to these principles, OSDR facilitates the secure and responsible reuse of valuable research data, thereby advancing scientific discovery while maintaining the trust, ethics and rights of data contributors (22,23).

Researchers can explore all OSDR data via the OSDR search functions, identifying studies and their access control levels. Restricted files are indicated by a lock icon and a 'Request Access to Restricted Files' button (Figure 5). To access restricted data, investigators must submit a controlled access data access request form, including IRB documentation, and agree to the data user agreement. The request is then verified by the OSDR Curation Team and reviewed by a Data Access Committee led by the original data submitter. Approved requests grant data access under specified terms, and investigators must sign and comply with these terms and provide progress updates. Upon project completion, all downloaded data must be deleted, and proper citations and acknowledgments are required.

Figure 5.



Open in a new tab

Data access request (DAR) process for OSDR controlled data. The process begins by clicking the 'Request Access to Restricted Files' button on datasets requiring controlled access (e.g. human sequence data). Requesters must submit a data request package with IRB documentation and agree to the Data User Agreement, which is reviewed by the Data Access Committee. Upon approval, data access is granted under the specified terms. Users are required to delete downloaded data upon project completion.

For example, users can request controlled data from the recent civilian astronaut mission Inspiration 4 (10,24). With this process, one can request raw data generated from whole blood and urine samples, and skin biopsies from the crew, collected at different time points before, during, and after spaceflight (i.e. OSD-569, OSD-570, OSD-571, OSD-574, OSD-575, OSD-656, OSD-687). This process facilitates secure data sharing, more diverse and inclusive access to spaceflight research, and a unique opportunity for the scientific community to unlock profound insights into the consequences of space travel on the human body without revealing sensitive data.

Enhancing usability and access

Data generation

The opportunities for biological experimentation in space are constrained by high costs, limited crew time, and limited availability of space aboard spacecraft and habitats, which restricts the number of spaceflight studies and their sample sizes. Therefore, to maximize scientific discoveries from spaceflight missions, NASA GeneLab established the SPL, which processes preserved leftover samples to generate new data or augment existing studies on OSDR. Tissue samples provided by the NASA Biological Institutional Scientific Collection (NBISC) and the Biospecimen Sharing Program from experiments conducted on the ISS or from legacy missions (e.g. Space Shuttle-era) are processed by SPL. The resulting data, complete with detailed metadata including a description of the origin mission, are promptly made available on OSDR for public scientific analysis, enhancing data accessibility and interpretability.

The SPL shares publicly accessible standard operating procedures (SOP) for various processes including tissue storage, cutting, homogenization, nucleotide extraction, library preparation, and sequencing (https://www.nasa.gov/osdr-sequencing-and-sample-processing-standard-operating-procedures/), developed in collaboration with the International Standards for Space Omics Processing consortium (https://issop.astrobotany.com/) (25). The use of consistent SOPs for data generation across distinct laboratories ensures harmonized DNA and RNA sequencing data from a wide range of biological samples, such as mammalian and plant tissues, cells and microbes.

Currently, OSDR hosts over 75 OSD studies produced by GeneLab's SPL, including more than 100 ‘omics datasets from experiments involving rodent, plant, and bacterial samples. These datasets have been instrumental in 31 peer-reviewed scientific publications, thereby significantly enriching our understanding of spaceflight's impacts on terrestrial biology (15,26–29). Additionally, GeneLab provides SPL services to NASA-funded PIs, allowing them to benefit from GeneLab's expertise in processing space-flown tissues (https://osdr.nasa.gov/bio/forms/genelab-sequencing-services.html). This simplifies data submission to OSDR, and the resulting data harmonization enables more accurate comparisons between studies.

Metadata and community consensus standards enables interoperability and reusability

Within OSDR, NASA GeneLab provides assay metadata standards and templates for PIs to follow when submitting through BDME for over 40 ‘omics assays (Table 1). The metadata for these assays follows the investigation, sample, assay (ISA)-tab format, a slightly altered version to meet the needs of OSDR (19,20) and uses standardized, controlled ontologies to ensure data is maximally FAIR. In 2022, ALSDA began curating non-omics datasets (e.g. physiological and phenotypic) using a similar approach.

Table 1.

OSDR assay metadata configurations[1]

'Omics assay types	Phenotypic-physiological assay types
AmpliconSequencing (ITS, 18S, 16S) (31)	ATPase activity via Spectrophotometry (32)
Whole Genome Bisulfite Sequencing	Rodent Gait (33)
Reduced Representation Bisulfite Sequencing	Open Field (34)
Whole Transcriptome Bisulfite Sequencing	Object in Place (35)
Enzymatic Methylation Sequencing	Radial Arm Water Maze (36)
Whole Genome Sequencing (37,38), MINISEQE Guidelines https://www.fged.org/projects/minseqe/)	Unconstrained Cognitive Flexibility (39)
Targeted DNA Sequencing	Novel Object Recognition (40)
Whole Genome Shotgun Metagenomics Sequencing (41)	Elevated Plus Maze (42,43)
Metatranscriptomics Sequencing	Attentional Set Shifting (44,45)
NucleotideMicroarray(DNA, RNA, miRNA, SNP) (46)	Barnes Maze (47)
Protein Microarray (46)	3-Chamber Social Isolation (48)
Bulk RNA Sequencing (49), MINISEQE Guidelines https://www.fged.org/projects/minseqe/ (49)	Drosophila Climbing Locomotion (50)
Targeted RNA Sequencing	Drosophila in-Flight Behavior (51)
Single Cell/Nuclei RNA Sequencing (52)	Operant Response (53)
microRNA Sequencing	Novel Object Location (54)
Single Cell/Nuclei ATAC Sequencing	Balance Beam (55)
Spatial Transcriptomics	Deacon Score (56)
Nanopore DNA Sequencing	3pt/4pt/Torsion Mechanical Tests (57)
Nanopore Targeted DNA Sequencing	Micro-computed Tomography (58)
Nanopore Direct RNA Sequencing	Dual-Energy X-Ray Absorptiometry (59)

‘Omics assay types	Phenotypic-physiological assay types
Targeted Circular Consensus DNA Sequencing	Peripheral Quantitative Computed Tomography (60)
Proteomic Mass Spectrometry (61)	Calcium Uptake via Spectrofluorometry (62)
Metabolomic Mass Spectrometry	Echocardiogram Ultrasound (63)
Metabolite NMR Spectroscopy	Estrous Cycle Cytology (64)
Chromatin Immunoprecipitation Sequencing	Flow Cytometry (65)
Chromatin Immunocleavage Sequencing	Genome Wide Association Study (66)
Methylated DNA Immunoprecipitation Sequencing	Histomorphometry (67)
Enzyme-Linked ImmunoSorbent Assay (ELISA)	Histology
	Light Microscopy
	Immunohistochemistry
	Immunostaining
	Fluorescence Microscopy
	Confocal Microscopy
	Transmission Electron Microscopy (68)
	Scanning Electron Microscopy (69)
	Morphometric Photography
	Muscle Contractility (70)
	Muscle Force via Force Plate (71)
	Muscle Force via Force Transducer (72)
	Ophthalmological Ultrasound (73)
	Optical Coherence Tomography (74)
	Western Blot (75)
	Magnetic Resonance Imaging (76,77)

Open in a new tab

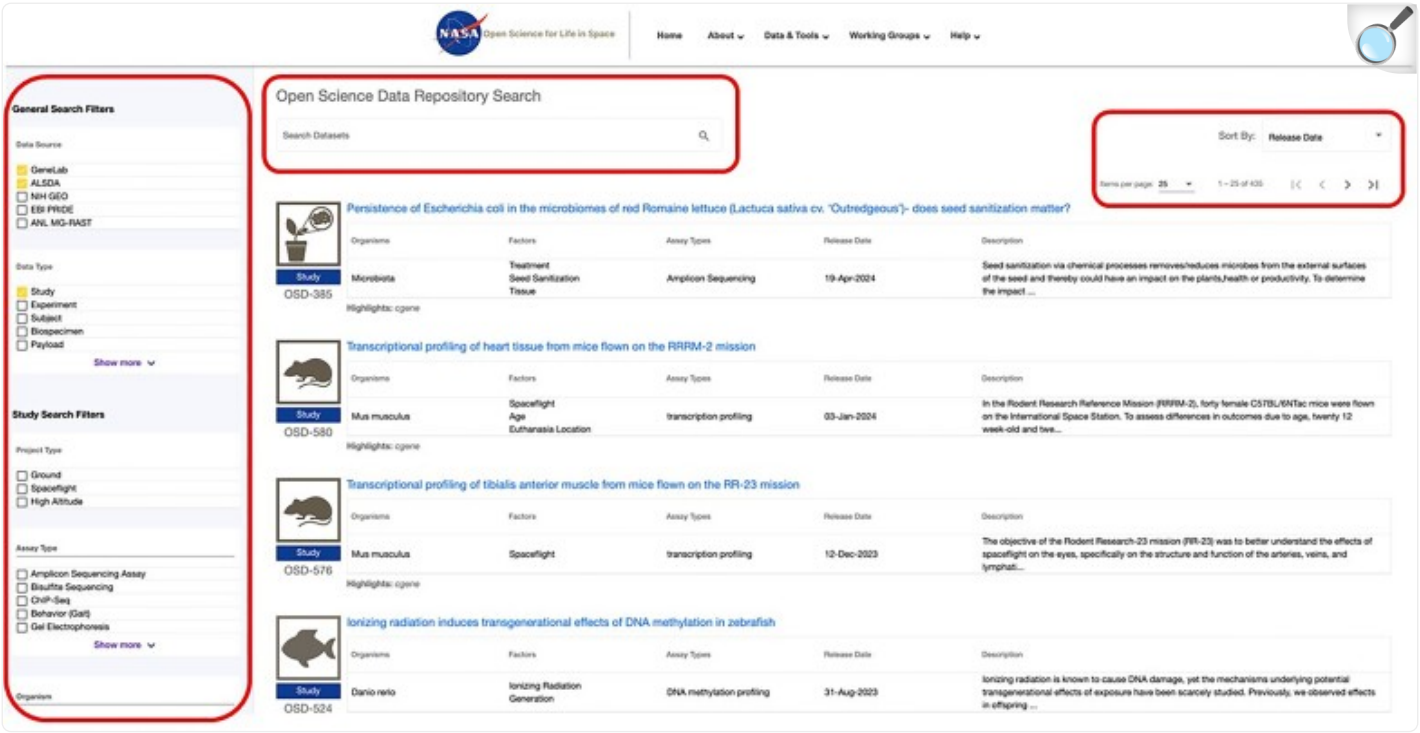
[1]List of standardized ‘omics and physiological-phenotypic assay types available through OSDR’s BDME. Assay standards were determined using established standards in the literature and through feedback and consensus discussions from AWG SMEs.

As of October 2024, 61 ALSDA assay metadata configurations have been created and implemented on OSDR. As more data types or new data types are submitted to OSDR, more are continuously being developed. These were established through virtual meetings with ∼150 subject matter experts (SMEs) and asynchronous documentation feedback within the ALSDA AWG. Each assay metadata standard was reviewed and discussed by at least three SMEs, with some of the more common assays reviewed by up to 10 experts. All OSDR assay metadata configurations with references are openly available for public review and used as metadata templates for OSDR submission (https://github.com/nasa/OSDR_Data_Curation). Table 1 shows the majority of the newly standardized phenotypic-physiological assay types available in OSDR, as well as the ‘omics assays. The curation and maximally open accessibility of these phenotypic-physiological spaceflight data have already resulted in two publications stemming from users’ data mining and reusing the data (28,30).

Data repository search

OSDR is equipped with expanded search features (Figure 6) that enable users to sort through studies using keyword searches and study search filters, including project type (e.g. spaceflight, ground, or high altitude), assay type (e.g. RNA sequencing, bisulfite sequencing, histomorphometry, mass spectrometry, ultrasonography, etc.), organism (e.g. rodent, human, plant, bacteria, etc.), tissue (e.g. leaf, seedlings, cells, feces, liver, retina, blood, etc.) and factors (e.g. spaceflight, ionizing radiation, age, sex, ecotype, etc.). Users can also filter searches based on data type, including study, experiment, subject, biospecimen, payload, mission, hardware and vehicle. They can specify their OSDR search to include ‘omics (GeneLab), non-omics (ALSDA) or both using the data source search filter. In addition to searching within OSDR, the data source search filter now includes federated search capabilities, allowing users to search other repositories, such as NIH GEO (https://www.ncbi.nlm.nih.gov/gds), EBI PRIDE (https://www.ebi.ac.uk/pride/), ANL MG-RAST (https://www.mg-rast.org/) and soon JGI (https://data.jgi.doe.gov/) and NMDC (https://data.microbiomedata.org/). These expanded search features enhance findability by helping users identify relevant datasets across different databases.

Figure 6.



Open in a new tab

OSDR search features. The OSDR repository landing page is equipped with a search bar to search based on keywords or phrases (top middle). Users can refine their search based on search filters including data source (e.g. GeneLab, ALSDA or federated search in NIH GEO), data type (e.g. study, experiment and subject), project type (e.g. ground, spaceflight), assay type (e.g. bisulfite sequencing, behavior-gait, RNA-seq), organism (e.g. rodent, human and plant), tissue (e.g. liver, root and cells) and factor (e.g. spaceflight, ionizing radiation) (left). Search results are displayed and can be sorted by relevance, release date, accession or title (top right).

Expanded API and direct cloud access

The OSDR API has been enhanced to include new data types, enabling users to query the database for experiment, mission, hardware, vehicle, subject and biospecimen data (https://www.nasa.gov/reference/osdr-public-api/). In addition, users can utilize the API to complete full-text searches, data file retrieval and metadata retrieval associated with both ‘omics and physiological-phenotypic data. The API empowers users to programmatically access data across multiple missions, experiments and datasets, facilitating meta-analysis and easing access to download large data files.

As part of the NASA Science Mission Directorate (SMD) Open-Source Science Initiative, AWS and NASA have established a Space Act Agreement to explore optimal practices for discovering, accessing and utilizing high-value NASA science datasets. All public data in OSDR are now available on the Registry of Open Data on AWS (https://registry.opendata.aws/nasa-osdr/), enabling users to directly query and execute scripts across extensive datasets using AWS S3 command-line tools. The scientific community gains significantly from this initiative, which removes the time required to download large datasets and allows for analysis using AWS's computational resources directly.

Increasing engagement

Analysis working groups

As space-related biological, biomedical, and other life science data/metadata are precious national and global resources, OSDR serves not just as a data repository, but as a central community hub for global space life sciences/biomedical data, with scientific results helping enable life to thrive in space.

Beyond making space biology and space health data FAIR and maximally open, OSDR (and its precursor GeneLab) was designed to maximally engage scientific community members to ensure OSDR is in line and up to date with scientific best practices, enable networking within and beyond the space biology community, and to build trust with users. This was accomplished by the development and expansion of the OSDR AWGs, which have grown into an enormous community around OSDR (https://awg.osdr.space/). There are ~800 volunteer OSDR AWG members, who provide feedback on scientific standards for reuse (subject and assay metadata; data processing pipelines; dataset formats and uniform structures for machine-readability) and collaborate to mine and reuse OSDR data to conduct scientific analyses. OSDR has enabled 92 publications as of October 2024, many directly from AWG collaborations and most notably in 'The Biology of Spaceflight' Cell Press package in 2020 (https://www.cell.com/c/the-biology-of-spaceflight) and the 'Space Omics and Medical Atlas (SOMA) across orbits' Nature Portfolio package in 2024 (https://www.nature.com/immersive/d42859-024-00009-8/index.html).

The OSDR AWGs are open to anyone in the world to join and currently consist of members from ~40 different countries. Participants range in career level from high school students to undergraduate and graduate students to postdocs to PIs and established scientists in a variety of different sectors including citizen scientists, academia, government, non-profits, NGOs and commercial industry.

Greater access, inclusion and quality with standardized data processing

'Omics data in their raw form are largely inaccessible to non-bioinformaticians; therefore, GeneLab collaborates with the scientific community via the AWG to develop standard processing pipelines to generate and publish processed data

on OSDR. Unlike raw data, processed data have greater immediate value to diverse users with varying technical backgrounds and computational capabilities. Furthermore, standardizing processing workflows is essential to match the pace of raw data generation, ensure reproducibility and enable standardized processed data for comparison across datasets.

At the second AWG annual workshop in 2019, GeneLab and the AWGs achieved a milestone by establishing and publishing the first RNAseq consensus pipeline (78). Building on this achievement, GeneLab has since developed standard processing pipelines for a variety of 'omics assays across different genomic, transcriptomic and epigenetic technologies. All these pipelines undergo a rigorous review process by the AWGs, incorporating feedback before their publication and implementation. These pipelines are publicly available on GitHub for broader access and use by the scientific community (https://github.com/nasa/GeneLab_Data_Processing). GeneLab pipelines are wrapped into Nextflow workflows (79) to automate and accelerate processing of these 'omics data types. In addition to core data processing, workflows also include raw data staging and a robust verification and validation program to identify errors in real-time, halt additional downstream computation and preserve computational resources.

To enhance reproducibility of GeneLab-generated data and expand access to GeneLab workflows for users without sufficient computing resources or the knowledge to install and run software, GeneLab partnered with the Department of Energy through a memorandum of understanding to create the NASA EDGE Bioinformatics platform (https://nasa.edgebioinformatics.org/home). NASA EDGE (Empowering the Development of Genomics Expertise; [(80)]) Bioinformatics is an open-access web-based platform that utilizes shared compute resources to run the GeneLab standardized bioinformatics workflows, which eliminates the need for researchers to have their own high-performance computing cluster. The web-based platform makes complicated biological analyses incredibly easy to perform, thus expanding the reach of these analyses to bioinformatics novices, students and even citizen scientists enabling them to contribute to scientific discoveries and progress.

Visualization tools enable knowledge discovery

Standardized processing pipelines in OSDR convert large volumes of raw data into tabular data that are easily interpretable, even for those without bioinformatics expertise, assuming the availability of suitable tools for data visualization. GeneLab recognized this need years ago by initiating the Visualization Platform, enabling users to delve into individual transcriptomic datasets. Building on this foundation, OSDR has introduced the Multi-Study Visualization Platform, enhancing functionality by allowing users to simultaneously compare data across multiple studies (https://visualization.genelab.nasa.gov/data/). Today, OSDR offers a comprehensive suite of visualization tools tailored for diverse research needs and designed to help users understand how environmental factors may affect OSDR biological data (Figure 7).

Figure 7.

Open in a new tab

OSDR Visualization Tools. Top left: Multi-study data visualization dashboard showing interactive processed transcriptomic data search filters, principal component analysis plot, differential gene expression table and normalized enrichment score plot. Top right: RadLab dashboard showing interactive radiation detection instrument selection, time-series plots and detector comparison plots. Bottom: EDA dashboard showing interactive temperature, humidity and CO_2 plots, mission milestones and multiple mission comparison plots.

The Multi-Study Visualization Platform enables the visualization of multiple RNA sequencing datasets simultaneously. Users can navigate through OSDR's comprehensive dataset repository, applying filters such as experimental factors, assay technology types, organisms and tissue types to pinpoint specific studies. Once users select samples from one or more RNA sequencing datasets to combine, they have the flexibility to normalize the combined data using the DESeq2 Median of Ratios method, followed by performing differential expression analysis. Additional selections can be made after examining both normalized and unnormalized PCA plots. The visualization interface includes interactive tools like pair plots, volcano plots, ideograms, gene set enrichment analysis and differential gene expression charts, all designed to enhance the user's ability to explore and analyze the data effectively.

Accessing environmental telemetry data of spacecraft relevant to the biological payloads available in OSDR has always been challenging due to data fragmentation and unclear data governance across space agencies and databases. To address this issue, OSDR introduced the Environmental Data Application (EDA; https://visualization.osdr.nasa.gov/eda), which transforms data streams from spacecraft into a time series format and reduces confounding factors among independent variables. As of October 2024, EDA focuses on telemetry and radiation data from Rodent Research missions aboard the ISS, encompassing temperature, relative humidity, CO_2 levels and radiation dose. The application enables users to analyze data from individual missions, compare two missions and download both summary and detailed data tables. Plans to expand EDA's capabilities to include additional biological payloads (human, plant, etc.) are contingent on receiving detailed data streams of information from the specific hardware and payloads involved.

In response to difficulties in accessing radiation data, OSDR supported the creation of RadLab, a database integrating radiation dosimetry from multiple space agencies including NASA, ESA, the German and Italian Space Agencies and the Bulgarian Academy of Sciences. RadLab features a visualization portal (https://visualization.osdr.nasa.gov/radlab) and an API (https://visualization.osdr.nasa.gov/radlab/api) that facilitates data queries by craft, sensor type, time range and radiation type (e.g. galactic cosmic rays, solar particle events, the contribution of the South Atlantic Anomaly), enhancing the accessibility and usability of these data. The API provides programmatic access for use in computational pipelines and GUIs for data visualization and exploration, making these data FAIR. OSDR is actively enhancing its integration with RadLab to provide average absorbed radiation doses for any payload, by incorporating these metadata directly into the repository. Additionally, radiation dose time series data are being directly fed into EDA to streamline access and analysis.

Training the next generation of space biologists and bioinformaticians

To sensitize the scientific community to space biology and increase the number of scientists who understand and utilize OSDR data and resources in 2021, OSDR introduced the GL4U training program (GeneLab for Colleges and Universities). GL4U leverages the consensus 'omics processing pipelines, developed by GeneLab and the AWGs, to teach students about space biology and bioinformatics via a series of lectures and hands-on training through Jupyter notebooks. The program offers annual 1–2-week long bootcamps for both students and educators (https://www.nasa.gov/reference/osdr-educator-resources/). During the bootcamps, participants complete the GL4U Introduction module set, which includes an introduction to NASA SMD, the NASA Space Biology program, OSDR and GeneLab, and hands-on training in basic coding in Unix and R. Participants then complete a GL4U 'omics-specific module set, which includes lectures and hands-on training in experimental design and data processing using the GeneLab standard pipeline. As of October, 2024, GL4U 'omics-specific module sets have been developed for RNA Sequencing and Amplicon Sequencing. To date, 14 educators and 76 students from 7 minority serving institutions have successfully completed a GL4U bootcamp. Pre- versus post-bootcamp surveys revealed a 115% increase, on average, in both participant understanding of 'omics data processing and in familiarity with NASA Space Biology resources, showing the overwhelming success of these bootcamps. GeneLab is expanding GL4U into a series of open-access on-

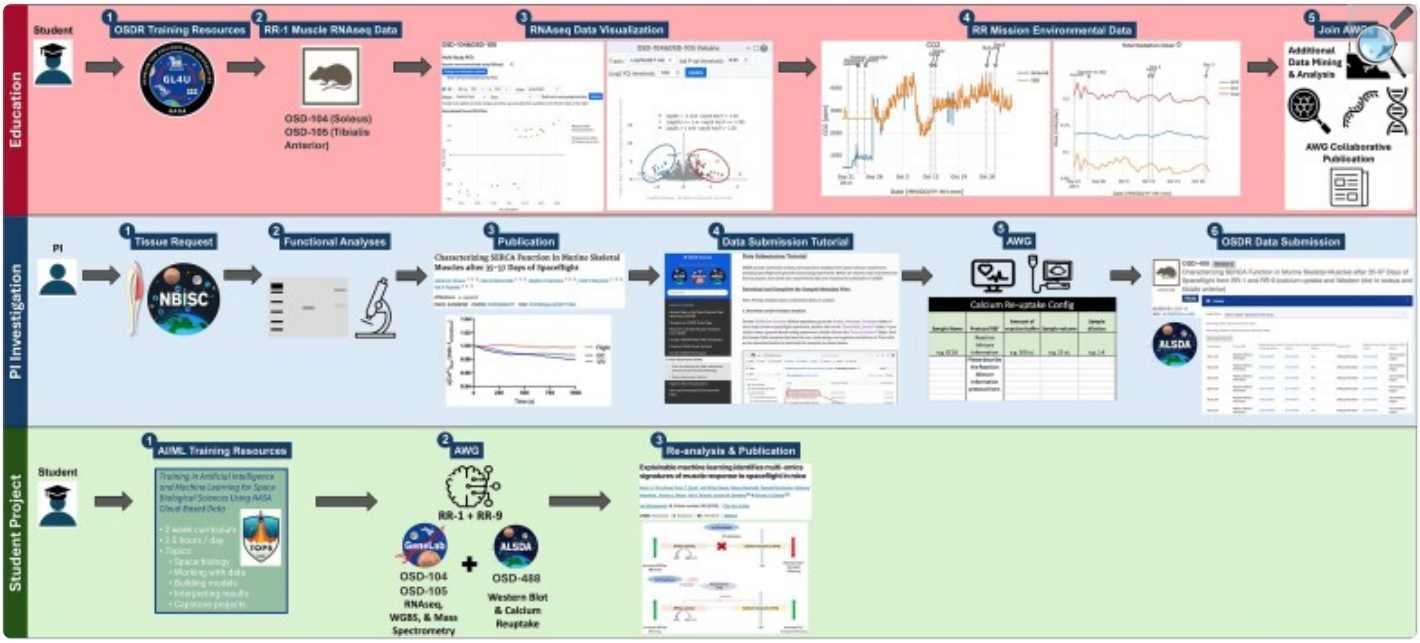
demand training modules.

NASA's SMD initiated the Transform to Open Science (TOPS) program in response to the 2022 White House mandate for open science for federal research, launching its operations in 2023, the Year of Open Science (18). The goal of TOPS is to rapidly transform agencies, organizations, and communities to an inclusive culture of open science in part by providing open-source training centered around open science. As part of this initiative, scientists from OSDR and the AWG were awarded a TOPS-T ScienceCore grant to create a training course titled: 'Training in Artificial Intelligence and Machine Learning for Space Biological Sciences Using NASA Cloud-Based Data'. This course, available on the Canvas platform (enroll here: https://canvas.instructure.com/enroll/8JYKD7), offers artificial intelligence and machine learning (AI/ML) training that uses space biology/health data, including OSDR datasets, to teach participants about space biology, model building, data handling and advanced modeling techniques. Students are shown how to interpret results through hands-on practice with real-world use cases from OSDR, using two large AI benchmark datasets released by OSDR for BPS on the Registry of Open Data on AWS (i.e. https://registry.opendata.aws/bps_microscopy/ and https://registry.opendata.aws/bps_rnaseq/). Originally offered synchronously through a live virtual course, the training is now also available asynchronously, allowing participants to engage with a rich array of data types such as phenotypic-physiological, biomedical, imaging and 'omics data, all hosted on OSDR with extensive, standardized metadata and processed data files. This curriculum paves the way for new scientific discoveries by leveraging AI/ML to mine the comprehensive data resources available through OSDR.

User roadmap for open science involvement

All data, metadata, resources and trainings available in the OSDR ecosystem are designed to help lower the bar for entry for users worldwide to engage and contribute scientifically to space exploration. Here, we outline three potential journeys of users to leverage different aspects of the OSDR ecosystem (Figure 8).

Figure 8.



Open in a new tab

User roadmap. Part 1: Student journey through OSDR training resources and visualization platforms. Part 2: PI journey through NBISC tissue request. Part 3: Student journey through TOPS-T AI/ML training resources and AWG collaboration.

In the first journey, a student may leverage the training and visualization aspects of the OSDR ecosystem by completing GL4U RNAseq bootcamp, gaining hands-on experience in space biology 'omics data and bioinformatics. Equipped with this knowledge, the student now has the tools to download and analyze data from OSDR and identifies OSD-104 and OSD-105 as studies with muscle multi-omics data from Rodent Research-1 (RR-1). The student uses the Multi-Study Data Visualization portal to combine both datasets, generating a PCA plot and a volcano plot from the RNAseq data and examining the differentially expressed genes between flight and ground control samples. The student notes that the PCA plot shows a strong bias toward tissue difference dominating the impact of spaceflight versus ground control. The student uses the EDA to visualize environmental variables during the RR-1 mission to investigate whether there are any environmental impacts contributing to the 'omics signals. To expand on this analysis, the student joins an AWG and collaborates with other researchers to explore the epigenomics and proteomics data available in these studies. This collaboration leads to the generation of new data that culminates in a scientific publication.

In the second journey, a PI with an established lab requests RR-1 and RR-9 muscle tissues from NBISC and performs

functional analysis in the lab. While publishing the findings, the PI submits data to OSDR using the data submission tutorial (https://osdr-tutorials.readthedocs.io/en/latest/) and submits the data into OSDR. One of the functional data types does not yet have an assay metadata configuration nor data format standard in OSDR, so the PI joins the AWG and helps create a new assay standard as part of a consensus with other AWG SME members with expertise on that assay. An example of such a workflow is Braun et al., 2021 (81), who requested NBISC samples, then analyzed muscle calcium uptake, helped create the OSDR assay configuration for such measurement and submitted the data to OSDR (OSD-488).

Finally, in the third journey, a student interested in using AI/ML approaches for space biology research enrolls in the asynchronous TOPS-T AI/ML for space biology training course. Enriched with new skills, the student proposes a project to the AWG to develop a predictive model using both RR-1 and RR-9 omics and functional data available on OSDR, leveraging the power of AI/ML for heterogeneous data types. The team re-analyzes the OSDR datasets and publishes their updated results, contributing valuable insights to the field. An example of a similar study is Li et al., 2023 (28), who analyzed 'omics and functional-physiological muscle data from OSDR using a ML workflow.

Images within this figure are adapted from: Braun J. et al. Characterizing SERCA function in murine skeletal muscles after 35–37 days of spaceflight. International Journal of Molecular Sciences, 2021, licensed under CC BY 4.0; Li K. et al. Explainable ML identifies multi-omics signatures of muscle response to spaceflight in mice. NPJ Microgravity, 2023, licensed under CC BY 4.0.

Discussion and future directions

The successful expansion of human life into deep space hinges on profound advancements in our understanding of space biology and the critical biomedical responses and health systems developed for long-duration and long-distance spaceflight missions. Centralizing and standardizing data, alongside the strategic reuse of tissues, are pivotal for maximizing the value extracted from every piece of available data and tissue. However, a meta-analysis of the field of space medicine and space life sciences by Bellomo et al. reviewed 1215 open-access articles and found that only a third shared their data. Analysis code was almost never shared, and practically no researchers used registered assay methodology standards (82). Especially as humanity nears returning to the Moon, this finding underscores the urgent need for a paradigm shift towards open science in space biology and aerospace medicine, an essential shift required to develop health countermeasures and conduct space biological research.

Responding to this need and leveraging the pioneering work of GeneLab (8), OSDR has rapidly expanded as a public resource to foster extensive engagement in spaceflight science, discovery, research, and development. This expansion has attracted tens of thousands of participants globally, especially via the AWG, and makes all relevant life sciences data from space experiments maximally open access and of high quality. Furthermore, ESA has recently adopted OSDR as its default portal for all life sciences ‘omics data, marking a significant international endorsement of OSDR standards.

This collaboration is poised to augment OSDR with hundreds of terabytes of ‘omics data soon, enhancing the repository's utility and scope.

Commercial astronauts are also actively participating in the space biology ecosystem by securely sharing their sensitive data through the OSDR DAR portal (10,18). Such practices not only show a shared commitment to centralizing valuable datasets between the scientific community and commercial space sectors but also enhance our ability to compare human, animal, microbial, plant and other data to deepen our understanding of space biology.

Acknowledging the limitations of traditional statistical methods (83), OSDR is also broadening its analytical range to include AI/ML technologies (4,84). In partnership with researchers from UC San Francisco, OSDR is integrating a knowledge graph analysis tool known as 'SPOKE' (Scalable Precision Medicine Open Knowledge Engine) into its visualization-analysis portal. This integration aims to enable users to identify symptoms, diseases, and potential therapeutic targets by mining and comparing datasets, enhancing the utility of OSDR for advanced data-driven discoveries (85–87).

OSDR is continuously refining its systems to provide better access and improved tools. The OSDR API will continue evolving to better serve users, with plans to develop more visualization tools for data analytics, including single-cell sequencing, metagenomics and physio-phenotypic-bioimaging data. AI/ML enhancements will streamline data submission, and an OSDR chatbot is being developed to help users discover and interpret datasets more easily.

To conclude, as humanity has entered the ‘Second Space Age’ (88), a central goal for the future of OSDR has evolved from FAIR data and metadata to the FAIREST principles, standing for more Engagement, more Social connections, and more Trust in its quality and stewardship (89). As the commercial spaceflight industry expands, OSDR is poised to facilitate broader public access to life science and biomedical data. As a leading force in open science, OSDR empowers the global community to actively engage in space exploration and discovery, aligning directly with NASA’s ambitious goals for the ‘Moon to Mars’ (90) and Artemis campaigns.

Acknowledgements

We wish to thank the scientific community including data submitters, users and advisors for their longstanding support of OSDR. We thank all OSDR AWG members for their years of feedback, support, and acts of scientific discovery in support of spaceflight life sciences. Numbering over 800 people, the robustness and successes of OSDR reflect this community of scientists, students and citizen scientists. Much of the activity and excitement around spaceflight life sciences and health would not be possible without the hard work of this community. Special thanks to the commercial crew, commercial researchers and international space agencies who collaborate and leverage OSDR for their datasets. We would like to thank all members of the OSDR team for their dedication, support and passion.

Contributor Information

Samrawit G Gebre, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Ryan T Scott, KBR, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Amanda M Saravia-Butler, KBR, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Danielle K Lopez, KBR, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Lauren M Sanders, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Sylvain V Costes, Space Biosciences Division, NASA Ames Research Center, Moffett Field, CA 94035, USA.

Data availability

All data are available on the NASA Open Science Data Repository: https://www.nasa.gov/osdr/ .

Funding

NASA Space Biology Program, part of the NASA Biological and Physical Sciences Division within the NASA Science Mission Directorate; NASA Human Research Program. Funding for open access: NASA.

Conflict of interest statement. None declared.

References

1. Afshinnekoo E., Scott R.T., MacKay M.J., Pariset E., Cekanaviciute E., Barker R., Gilroy S., Hassane D., Smith S.M., Zwart S.R.et al.. Fundamental biological features of spaceflight: advancing the field to enable deep-space exploration. Cell. 2020; 183:1162–1184. [DOI] [PMC free article] [PubMed] [Google Scholar]

2. Blaber E., Boothby T., Carr C.E., Craig Everroad R., Foster J., Galazka J., Lee J.A., Lera M., Ricco A., Sanders L.et al.. Space biology beyond LEO instrumentation & science series science working group 2022

annual report. NASA Techn. Rep. Server. 2023; 20230008417. [Google Scholar]

3. National Academies of Sciences, Engineering, and Medicine; Division on Engineering and Physical Sciences; Aeronautics and Space Engineering Board; Space Studies Board; Committee on Biological and Physical Sciences Research in Space 2023–2032 Thriving in Space: Ensuring the Future of Biological and Physical Sciences Research: A Decadal Survey for 2023-2032. 2023; Washington (DC).National Academies Press(US). [PubMed] [Google Scholar]

4. Scott R.T., Sanders L.M., Antonsen E.L., Hastings J.J.A., Park S.-M., Mackintosh G., Reynolds R.J., Hoarfrost A.L., Sawyer A., Greene C.S.et al.. Biomonitoring and precision health in deep space supported by artificial intelligence. Nat. Mach. Intell. 2023; 5:196–207. [Google Scholar]

5. Wilkinson M.D., Dumontier M., Aalbersberg I.J.J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J.-W., da Silva Santos L.B., Bourne P.E.et al.. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data. 2016; 3:160018. [DOI] [PMC free article] [PubMed] [Google Scholar]

6. Stall S., Yarmey L., Cutcher-Gershenfeld J., Hanson B., Lehnert K., Nosek B., Parsons M., Robinson E., Wyborn L.. Make scientific data FAIR. Nature. 2019; 570:27–29. [DOI] [PubMed] [Google Scholar]

7. Higman R., Bangert D., Jones S.. Three camps, one destination: the intersections of research data management, FAIR and Open. Insights Imaging. 2019; 32:1–9. [Google Scholar]

8. Berrios D.C., Galazka J., Grigorev K., Gebre S., Costes S.V.. NASA GeneLab: interfaces for the exploration of space omics data. Nucleic Acids Res. 2021; 49:D1515–D1522. [DOI] [PMC free article] [PubMed] [Google Scholar]

9. da Silveira W.A., Fazelinia H., Rosenthal S.B., Laiakis E.C., Kim M.S., Meydan C., Kidane Y., Rathi K.S., Smith S.M., Stear B.et al.. Comprehensive multi-omics analysis reveals mitochondrial stress as a central biological hub for spaceflight impact. Cell. 2020; 183:1185–1201. [DOI] [PMC free article] [PubMed] [Google Scholar]

10. Sanders L.M., Grigorev K.A., Scott R.T., Saravia-Butler A.M., Polo S.-H.L., Gilbert R., Overbey E.G., Kim J., Mason C.E., Costes S.V.. Inspiration4 data access through the NASA Open Science Data Repository. NPJ Microgravity. 2024; 10:56. [DOI] [PMC free article] [PubMed] [Google Scholar]

11. Siew K., Nestler K.A., Nelson C., D'Ambrosio V., Zhong C., Li Z., Grillo A., Wan E.R., Patel V., Overbey E.et al.. Cosmic kidney disease: an integrated pan-omic, physiological and morphological study into spaceflight-induced renal dysfunction. Nat. Commun. 2024; 15:4923. [DOI] [PMC free article] [PubMed] [Google Scholar]

12. Mathyk B.A., Tabetah M., Karim R., Zaksas V., Kim J., Anu R.I., Muratani M., Tasoula A., Singh R.S.,

Chen Y.-K.et al.. Spaceflight induces changes in gene expression profiles linked to insulin and estrogen. Commun. Biol. 2024; 7:692. [DOI] [PMC free article] [PubMed] [Google Scholar]

13. Beheshti A., Cekanaviciute E., Smith D.J., Costes S.V.. Global transcriptomic analysis suggests carbon dioxide as an environmental stressor in spaceflight: a systems biology GeneLab case study. Sci. Rep. 2018; 8:4191. [DOI] [PMC free article] [PubMed] [Google Scholar]

14. Cope H., Elsborg J., Demharter S., McDonald J.T., Wernecke C., Parthasarathy H., Unadkat H., Chatrathi M., Claudio J., Reinsch S.et al.. Transcriptomics analysis reveals molecular alterations underpinning spaceflight dermatology. Commun. Med. 2024; 4:106. [DOI] [PMC free article] [PubMed] [Google Scholar]

15. Beheshti A., Chakravarty K., Fogle H., Fazelinia H., Silveira W.A.d., Boyko V., Polo S.-H.L., Saravia-Butler A.M., Hardiman G., Taylor D.et al.. Multi-omics analysis of multiple missions to space reveal a theme of lipid dysregulation in mouse liver. Sci. Rep. 2019; 9:19195. [DOI] [PMC free article] [PubMed] [Google Scholar]

16. Jiang P., Green S.J., Chlipala G.E., Turek F.W., Vitaterna M.H.. Reproducible changes in the gut microbiome suggest a shift in microbial and host metabolism during spaceflight. Microbiome. 2019; 7:113. [DOI] [PMC free article] [PubMed] [Google Scholar]

17. Scott R.T., Grigorev K., Mackintosh G., Gebre S.G., Mason C.E., Del Alto M.E., Costes S.V.. Advancing the Integration of Biosciences Data Sharing to Further Enable Space Exploration. Cell Rep. 2020; 33:108441. [DOI] [PubMed] [Google Scholar]

18. Costes S., Gentemann C., Platts S.H., Carnell L.A.. Biological horizons: pioneering open science in the cosmos. Nat. Commun. 2024; 15:4780. [DOI] [PMC free article] [PubMed] [Google Scholar]

19. Sansone S.-A., Rocca-Serra P., Brandizi M., Brazma A., Field D., Fostel J., Garrow A.G., Gilbert J., Goodsaid F., Hardy N.et al.. The first RSBI (ISA-TAB) Workshop: ‘Can a Simple Format Work for Complex Studies?’. OMICS. 2008; 12:143–149. [DOI] [PubMed] [Google Scholar]

20. González-Beltrán A., Maguire E., Sansone S.-A., Rocca-Serra P.. linkedISA: semantic representation of ISA-Tab experimental metadata. BMC Bioinf. 2014; 15:1–15. [DOI] [PMC free article] [PubMed] [Google Scholar]

21. Tryka K.A., Hao L., Sturcke A., Jin Y., Wang Z.Y., Ziyabari L., Lee M., Popova N., Sharopova N., Kimura M.et al.. NCBI’s database of genotypes and phenotypes: dbGaP. Nucleic Acids Res. 2014; 42:D975–D979. [DOI] [PMC free article] [PubMed] [Google Scholar]

22. Rahimzadeh V., Fogarty J., Caulfield T., Auñón-Chancellor S., Borry P., Candia J., Cohen I.G., Covington

M., Lynch H.F., Greely H.T.et al.. Ethically cleared to launch?. Science. 2023; 381:1408–1411. [DOI] [PubMed] [Google Scholar]

23. Seylani A., Galsinh A.S., Tasoula A., I A.R., Camera A., Calleja-Agius J., Borg J., Goel C., Kim J., Clark K.B.et al.. Ethical considerations for the age of non-governmental space exploration. Nat. Commun. 2024; 15:4774. [DOI] [PMC free article] [PubMed] [Google Scholar]

24. Jones C.W., Overbey E.G., Lacombe J., Ecker A.J., Meydan C., Ryon K., Tierney B., Damle N., MacKay M., Afshin E.E.et al.. Molecular and physiological changes in the SpaceX Inspiration4 civilian crew. Nature. 2024; 632:1155–1164. [DOI] [PMC free article] [PubMed] [Google Scholar]

25. Rutter L., Barker R., Bezdan D., Cope H., Costes S.V., Degoricija L., Fisch K.M., Gabitto M.I., Gebre S., Giacomello S.et al.. A new era for space life science: international standards for space omics processing. Patterns (N Y). 2020; 1:100148. [DOI] [PMC free article] [PubMed] [Google Scholar]

26. Camera A., Tabetah M., Castañeda V., Kim J., Galsinh A.S., Haro-Vinueza A., Salinas I., Seylani A., Arif S., Das S.et al.. Aging and putative frailty biomarkers are altered by spaceflight. Sci. Rep. 2024; 14:13098. [DOI] [PMC free article] [PubMed] [Google Scholar]

27. Vitry G., Finch R., Mcstay G., Behesti A., Déjean S., Larose T., Wotring V., da Silveira W.A.. Muscle atrophy phenotype gene expression during spaceflight is linked to a metabolic crosstalk in both the liver and the muscle in mice. iScience. 2022; 25:105213. [DOI] [PMC free article] [PubMed] [Google Scholar]

28. Li K., Desai R., Scott R.T., Steele J.R., Machado M., Demharter S., Hoarfrost A., Braun J.L., Fajardo V.A., Sanders L.M.et al.. Explainable machine learning identifies multi-omics signatures of muscle response to spaceflight in mice. NPJ Microgravity. 2023; 9:90. [DOI] [PMC free article] [PubMed] [Google Scholar]

29. Sanders L.M., Chok H., Samson F., Acuna A.U., Polo S.-H.L., Boyko V., Chen Y.-C., Dinh M., Gebre S., Galazka J.M.et al.. Batch effect correction methods for NASA GeneLab transcriptomic datasets. Front. Astron. Space Sci. 2023; 10:1200132. [Google Scholar]

30. Reynolds R.J., Scott R.T., Turner R.T., Iwaniec U.T., Bouxsein M.L., Sanders L.M., Antonsen E.L.. Validating causal diagrams of human health risks for spaceflight: An example using bone data from rodents. Biomedicines. 2022; 10:2187. [DOI] [PMC free article] [PubMed] [Google Scholar]

31. Keenum I., Player R., Kralj J., Servetas S., Sussman M.D., Russell J.A., Stone J., Chandrapati S., Sozhamannan S.. Amplicon sequencing minimal information (ASqMI): quality and reporting guidelines for actionable calls in biodefense applications. J. AOAC Int. 2023; 106:1424–1430. [DOI] [PMC free article] [PubMed] [Google Scholar]

32. Rule C.S., Patrick M., Sandkvist M.. Measuring in vitro ATPase activity for enzymatic characterization. J. Vis. Exp. 2016; 114:e54305. [DOI] [PMC free article] [PubMed] [Google Scholar]

33. Dorman C.W., Krug H.E., Frizelle S.P., Funkenbusch S., Mahowald M.L.. A comparison of DigiGait™ and TreadScan™ imaging systems: assessment of pain using gait analysis in murine monoarthritis. J. Pain Res. 2014; 7:25–35. [DOI] [PMC free article] [PubMed] [Google Scholar]

34. Seibenhener M.L., Wooten M.C.. Use of the Open Field Maze to measure locomotor and anxiety-like behavior in mice. J. Vis. Exp. 2015; 96:e52434. [DOI] [PMC free article] [PubMed] [Google Scholar]

35. Huang H., Jiang N., Zhang Y.W., Lv J.W., Wang H.X., Lu C., Liu X.M., Lu G.H.. Gastrodia elata blume ameliorates circadian rhythm disorder-induced mice memory impairment. Life Sci. Space Res. 2021; 31:51–58. [DOI] [PubMed] [Google Scholar]

36. Penley S.C., Gaudet C.M., Threlkeld S.W.. Use of an eight-arm radial water maze to assess working and reference memory following neonatal brain injury. J. Vis. Exp. 2013; 82:e50940. [DOI] [PMC free article] [PubMed] [Google Scholar]

37. Field D., Garrity G., Gray T., Morrison N., Selengut J., Sterk P., Tatusova T., Thomson N., Allen M.J., Angiuoli S.V.et al.. The minimum information about a genome sequence (MIGS) specification. Nat. Biotechnol. 2008; 26:541–547. [DOI] [PMC free article] [PubMed] [Google Scholar]

38. Yilmaz P., Kottmann R., Field D., Knight R., Cole J.R., Amaral-Zettler L., Gilbert J.A., Karsch-Mizrachi I., Johnston A., Cochrane G.et al.. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat. Biotechnol. 2011; 29:415–420. [DOI] [PMC free article] [PubMed] [Google Scholar]

39. Britten R.A., Duncan V.D., Fesshaye A., Rudobeck E., Nelson G.A., Vlkolinsky R.. Altered cognitive flexibility and synaptic plasticity in the rat prefrontal cortex after exposure to low (≤15 cGy) doses of 28Si radiation. Radiat. Res. 2020; 193:223–235. [DOI] [PubMed] [Google Scholar]

40. Lueptow L.M. Novel object recognition test for the investigation of learning and memory in mice. J. Vis. Exp. 2017; 126:e55718. [DOI] [PMC free article] [PubMed] [Google Scholar]

41. Bowers R.M., Kyrpides N.C., Stepanauskas R., Harmon-Smith M., Doud D., Reddy T.B.K., Schulz F., Jarett J., Rivers A.R., Eloe-Fadrosh E.A.et al.. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 2017; 35:725–731. [DOI] [PMC free article] [PubMed] [Google Scholar]

42. Komada M., Takao K., Miyakawa T.. Elevated plus maze for mice. J. Vis. Exp. 2008; 22:e1088. [DOI] [PMC free article] [PubMed] [Google Scholar]

43. Walf A.A., Frye C.A.. The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. Nat. Protoc. 2007; 2:322–328. [DOI] [PMC free article] [PubMed] [Google Scholar]

44. Tait D.S., Bowman E.M., Neuwirth L.S., Brown V.J.. Assessment of intradimensional/extradimensional attentional set-shifting in rats. Neurosci. Biobehav. Rev. 2018; 89:72–84. [DOI] [PubMed] [Google Scholar]

45. Jewell J.S., Duncan V.D., Fesshaye A., Tondin A., Macadat E., Britten R.A.. Exposure to ≤15 cGy of 600 MeV/n 56Fe particles impairs rule acquisition but not long-term memory in the attentional set-shifting assay. Radiat. Res. 2018; 190:565–575. [DOI] [PubMed] [Google Scholar]

46. Brazma A., Hingamp P., Quackenbush J., Sherlock G., Spellman P., Stoeckert C., Aach J., Ansorge W., Ball C.A., Causton H.C.et al.. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat. Genet. 2001; 29:365–371. [DOI] [PubMed] [Google Scholar]

47. Barnes C.A., Jung M.W., McNaughton B.L., Korol D.L., Andreasson K., Worley P.F.. LTP saturation and spatial learning disruption: effects of task variables and saturation levels. J. Neurosci. 1994; 14:5793–5806. [DOI] [PMC free article] [PubMed] [Google Scholar]

48. Kaidanovich-Beilin O., Lipina T., Vukobradovic I., Roder J., Woodgett J.R.. Assessment of social interaction behaviors. J. Vis. Exp. 2011; 48:e2473. [DOI] [PMC free article] [PubMed] [Google Scholar]

49. Kolesnikov N., Hastings E., Keays M., Melnichuk O., Tang Y.A., Williams E., Dylag M., Kurbatova N., Brandizi M., Burdett T.et al.. ArrayExpress update–simplifying data submissions. Nucleic Acids Res. 2015; 43:D1113–D1116. [DOI] [PMC free article] [PubMed] [Google Scholar]

50. Le Bourg E., Lints F.A.. Hypergravity and aging in Drosophila melanogaster. 4. Climbing activity. Gerontology. 1992; 38:59–64. [DOI] [PubMed] [Google Scholar]

51. Mhatre S.D., Iyer J., Petereit J., Dolling-Boreham R.M., Tyryshkina A., Paul A.M., Gilbert R., Jensen M., Woolsey R.J., Anand S.et al.. Artificial gravity partially protects space-induced neurological deficits in Drosophila melanogaster. Cell Rep. 2022; 40:111279. [DOI] [PMC free article] [PubMed] [Google Scholar]

52. Füllgrabe A., George N., Green M., Nejad P., Aronow B., Fexova S.K., Fischer C., Freeberg M.A., Huerta L., Morrison N.et al.. Guidelines for reporting single-cell RNA-seq experiments. Nat. Biotechnol. 2020; 38:1384–1386. [DOI] [PMC free article] [PubMed] [Google Scholar]

53. Rabin B.M., Miller M.G., Larsen A., Spadafora C., Zolnerowich N.N., Dell'Acqua L.A., Shukitt-Hale B.. Effects of exposure to 12C and 4He particles on cognitive performance of intact and ovariectomized female rats. Life Sci. Space Res. 2019; 22:47–54. [DOI] [PubMed] [Google Scholar]

54. Denninger J.K., Smith B.M., Kirby E.D.. Novel object recognition and object location behavioral testing in mice on a budget. J. Vis. Exp. 2018; 141:e58593. [DOI] [PMC free article] [PubMed] [Google Scholar]

55. Feng Q., Ma Y., Mu S., Wu J., Chen S., Ouyang L., Lei W.. Specific reactions of different striatal neuron types in morphology induced by quinolinic acid in rats. PLoS One. 2014; 9:e91512. [DOI] [PMC free article] [PubMed] [Google Scholar]

56. Deacon R.M.J. Assessing nest building in mice. Nat. Protoc. 2006; 1:1117–1119. [DOI] [PubMed] [Google Scholar]

57. Jepsen K.J., Silva M.J., Vashishth D., Guo X.E., van der Meulen M.C.H.. Establishing biomechanical mechanisms in mouse models: practical guidelines for systematically evaluating phenotypic changes in the diaphyses of long bones. J. Bone Miner. Res. 2015; 30:951–966. [DOI] [PMC free article] [PubMed] [Google Scholar]

58. Bouxsein M.L., Boyd S.K., Christiansen B.A., Guldberg R.E., Jepsen K.J., Müller R.. Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. J. Bone Miner. Res. 2010; 25:1468–1486. [DOI] [PubMed] [Google Scholar]

59. Nagy T.R., Clair A.L.. Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice. Obes. Res. 2000; 8:392–398. [DOI] [PubMed] [Google Scholar]

60. Whittier D.E., Boyd S.K., Burghardt A.J., Paccou J., Ghasem-Zadeh A., Chapurlat R., Engelke K., Bouxsein M.L.. Guidelines for the assessment of bone density and microarchitecture in vivo using high-resolution peripheral quantitative computed tomography. Osteoporos. Int. 2020; 31:1607–1627. [DOI] [PMC free article] [PubMed] [Google Scholar]

61. Taylor C.F., Paton N.W., Lilley K.S., Binz P.-A., Julian R.K. Jr, Jones A.R., Zhu W., Apweiler R., Aebersold R., Deutsch E.W.et al.. The minimum information about a proteomics experiment (MIAPE). Nat. Biotechnol. 2007; 25:887–893. [DOI] [PubMed] [Google Scholar]

62. Tupling R., Green H.. Silver ions induce Ca2+ release from the SR in vitro by acting on the Ca2+ release channel and the Ca2+ pump. J. Appl. Physiol. 2002; 92:1603–1610. [DOI] [PubMed] [Google Scholar]

63. Wu J., Bu L., Gong H., Jiang G., Li L., Ma H., Zhou N., Lin L., Chen Z., Ye Y.et al.. Effects of heart rate and anesthetic timing on high-resolution echocardiographic assessment under isoflurane anesthesia in mice. J. Ultrasound Med. 2010; 29:1771–1778. [DOI] [PubMed] [Google Scholar]

64. McLean A.C., Valenzuela N., Fai S., Bennett S.A.L.. Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. J. Vis. Exp. 2012; 67:e4389.

[DOI] [PMC free article] [PubMed] [Google Scholar]

65. McKinnon K.M. Flow cytometry: an overview. Curr. Protoc. Immunol. 2018; 120:1–16. [DOI] [PMC free article] [PubMed] [Google Scholar]

66. Robinson J.R., Denny J.C., Roden D.M., Van Driest S.L.. Genome-wide and phenome-wide approaches to understand variable drug actions in electronic health records. Clin. Transl. Sci. 2018; 11:112–122. [DOI] [PMC free article] [PubMed] [Google Scholar]

67. Dempster D.W., Compston J.E., Drezner M.K., Glorieux F.H., Kanis J.A., Malluche H., Meunier P.J., Ott S.M., Recker R.R., Parfitt A.M.. Standardized nomenclature, symbols, and units for bone histomorphometry: a 2012 update of the report of the ASBMR Histomorphometry Nomenclature Committee. J. Bone Miner. Res. 2013; 28:2–17. [DOI] [PMC free article] [PubMed] [Google Scholar]

68. Taheri M.L., Stach E.A., Arslan I., Crozier P.A., Kabius B.C., LaGrange T., Minor A.M., Takeda S., Tanase M., Wagner J.B.et al.. Current status and future directions for in situ transmission electron microscopy. Ultramicroscopy. 2016; 170:86–95. [DOI] [PMC free article] [PubMed] [Google Scholar]

69. Golding C.G., Lamboo L.L., Beniac D.R., Booth T.F.. The scanning electron microscope in microbiology and diagnosis of infectious disease. Sci. Rep. 2016; 6:26516. [DOI] [PMC free article] [PubMed] [Google Scholar]

70. Gittings W., Bunda J., Stull J.T., Vandenboom R.. Interaction of posttetanic potentiation and the catchlike property in mouse skeletal muscle. Muscle Nerve. 2016; 54:308–316. [DOI] [PubMed] [Google Scholar]

71. Sobotka S., Mu L.. Characteristics of tetanic force produced by the sternomastoid muscle of the rat. J. Biomed. Biotechnol. 2010; 2010:194984. [DOI] [PMC free article] [PubMed] [Google Scholar]

72. Takeshita H., Yamamoto K., Nozato S., Inagaki T., Tsuchimochi H., Shirai M., Yamamoto R., Imaizumi Y., Hongyo K., Yokoyama S.et al.. Modified forelimb grip strength test detects aging-associated physiological decline in skeletal muscle function in male mice. Sci. Rep. 2017; 7:42323. [DOI] [PMC free article] [PubMed] [Google Scholar]

73. Rose L.T., Moshegov C.N.. Comparison of the Zeiss IOLMaster and applanation A-scan ultrasound: biometry for intraocular lens calculation. Clin. Exp. Ophthalmol. 2003; 31:121–124. [DOI] [PubMed] [Google Scholar]

74. Rahman W., Chen F.K., Yeoh J., Patel P., Tufail A., Da Cruz L.. Repeatability of manual subfoveal choroidal thickness measurements in healthy subjects using the technique of enhanced depth imaging optical coherence tomography. Invest. Ophthalmol. Vis. Sci. 2011; 52:2267–2271. [DOI] [PubMed] [Google Scholar]

75. Mahmood T., Yang P.-C.. Western blot: technique, theory, and trouble shooting. N. Am. J. Med. Sci. 2012; 4:429–434. [DOI] [PMC free article] [PubMed] [Google Scholar]

76. Grover V.P.B., Tognarelli J.M., Crossey M.M.E., Cox I.J., Taylor-Robinson S.D., McPhail M.J.W.. Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. J. Clin. Exp. Hepatol. 2015; 5:246–255. [DOI] [PMC free article] [PubMed] [Google Scholar]

77. Pautler R.G. Mouse MRI: concepts and applications in physiology. Physiology. 2004; 19:168–175. [DOI] [PubMed] [Google Scholar]

78. Overbey E.G., Saravia-Butler A.M., Zhang Z., Rathi K.S., Fogle H., da Silveira W.A., Barker R.J., Bass J.J., Beheshti A., Berrios D.C.et al.. NASA GeneLab RNA-seq consensus pipeline: standardized processing of short-read RNA-seq data. iScience. 2021; 24:102361. [DOI] [PMC free article] [PubMed] [Google Scholar]

79. Di Tommaso P., Chatzou M., Floden E.W., Barja P.P., Palumbo E., Notredame C.. Nextflow enables reproducible computational workflows. Nat. Biotechnol. 2017; 35:316–319. [DOI] [PubMed] [Google Scholar]

80. Li P.-E., Lo C.-C., Anderson J.J., Davenport K.W., Bishop-Lilly K.A., Xu Y., Ahmed S., Feng S., Mokashi V.P., Chain P.S.G.. Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform. Nucleic Acids Res. 2017; 45:67–80. [DOI] [PMC free article] [PubMed] [Google Scholar]

81. Braun J.L., Geromella M.S., Hamstra S.I., Messner H.N., Fajardo V.A.. Characterizing SERCA function in Murine skeletal muscles after 35-37 days of spaceflight. Int. J. Mol. Sci. 2021; 22:11764. [DOI] [PMC free article] [PubMed] [Google Scholar]

82. Bellomo R.K., Zavalis E.A., Ioannidis J.P.A.. Assessment of transparency indicators in space medicine. PLoS One. 2024; 19:e0300701. [DOI] [PMC free article] [PubMed] [Google Scholar]

83. Reynolds R.J., Shelhamer M.. Reynolds R.J. Research methods for the next 60 years of space exploration, BeyondLEO-Human Health Issues for Deep Space Exploration. Beyond LEO. 2020; London: IntechOpen; 1–7. [Google Scholar]

84. Sanders L.M., Scott R.T., Yang J.H., Qutub A.A., Garcia Martin H., Berrios D.C., Hastings J.J.A., Rask J., Mackintosh G., Hoarfrost A.L.et al.. Biological research and self-driving labs in deep space supported by artificial intelligence. Nat. Mach. Intell. 2023; 5:208–219. [Google Scholar]

85. Nelson C.A., Butte A.J., Baranzini S.E.. Integrating biomedical research and electronic health records to create knowledge-based biologically meaningful machine-readable embeddings. Nat. Commun. 2019;

10:3045. [DOI] [PMC free article] [PubMed] [Google Scholar]

86. Morris J.H., Soman K., Akbas R.E., Zhou X., Smith B., Meng E.C., Huang C.C., Cerono G., Schenk G., Rizk-Jackson A.et al.. The scalable precision medicine open knowledge engine (SPOKE): a massive knowledge graph of biomedical information. Bioinformatics. 2023; 39:btad080. [DOI] [PMC free article] [PubMed] [Google Scholar]

87. Nelson C.A., Acuna A.U., Paul A.M., Scott R.T., Butte A.J., Cekanaviciute E., Baranzini S.E., Costes S.V.. Knowledge network embedding of transcriptomic data from spaceflown mice uncovers signs and symptoms associated with terrestrial diseases. Life. 2021; 11:42. [DOI] [PMC free article] [PubMed] [Google Scholar]

88. Mason C.E., Green J., Adamopoulos K.I., Afshin E.E., Baechle J.J., Basner M., Bailey S.M., Bielski L., Borg J., Borg J.et al.. A second space age spanning omics, platforms, and medicine across orbits. Nature. 2024; 632:995–1008. [DOI] [PMC free article] [PubMed] [Google Scholar]

89. d'Aquin M., Kirstein F., Oliveira D., Schimmler S., Urbanek S.. FAIREST: a framework for assessing research repositories. Data Intell. 2023; 5:202–241. [Google Scholar]

90. Goodliff K.E., Merancy N.F., Bhakta S.S., Rucker M.A., Chai P.R.P., Ashurst T.E., Troutman P.A., Stromgren C.. Exploration Systems Development Mission Directorate (ESDMD) moon-to-mars architecture definition document. NASA Techn. Rep. Server. 2023; 20230002706. [Google Scholar]

Associated Data

This section collects any data citations, data availability statements, or supplementary materials included in this article.

Data Availability Statement

All data are available on the NASA Open Science Data Repository: https://www.nasa.gov/osdr/ .